SIGCOMM 2017 Topic Preview (Programmable Devices)

Rohan Gandhi
CMU

Programmable Devices

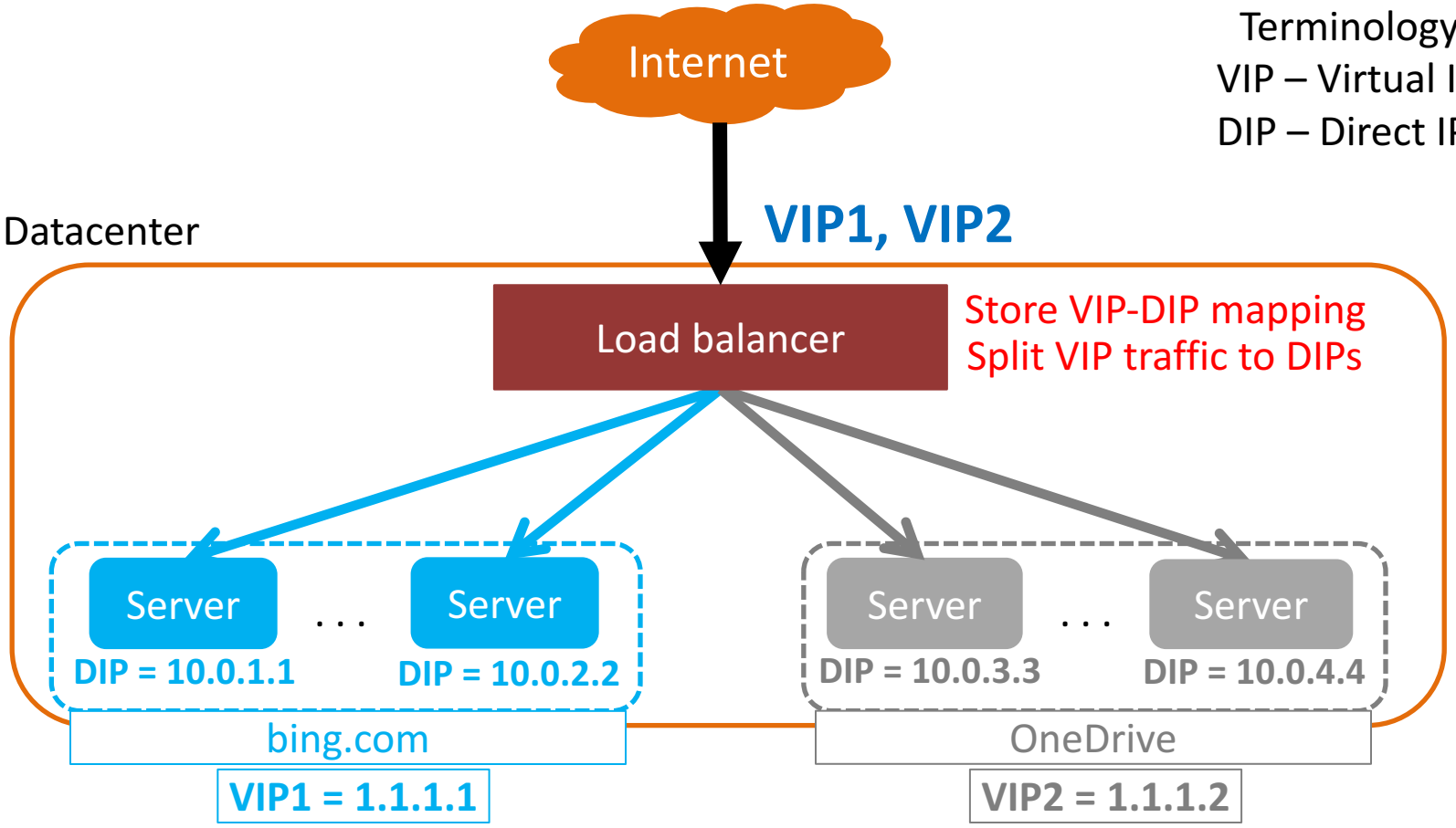
- Session-1 (11:15AM tomorrow), Centennial hall

1. DRMT: Disaggregated Programmable Switching
2. SilkRoad: Making Stateful Layer-4 Load Balancing Fast and Cheap Using Switching ASICs
3. Re-architecting datacenter networks and stacks for low latency and high performance

SilkRoad: Making Stateful Layer-4 Load Balancing Fast and Cheap Using Switching ASICs

Rui Miao (University of Southern California), Hongyi Zeng (Facebook), Changhoon Kim and Jeongkeun Lee (Barefoot Networks), and Minlan Yu (Yale University)

What Is A Layer-4 Load Balancer?



Load balancer provides high availability and scalability

Load Balancer Scale

Load balancer handles all Internet traffic + fraction of intra-DC traffic

	VIPs (Virtual IPs)	DIPs (Direct IPs)	Traffic
2013[1]	100K+	1 million+	40 Tbps+

[1]Ananta: Cloud Scale Load-balancing, SIGCOMM 2013

Prior Research Work

- Ananta (SIGCOMM 2013)
  - Software based LB at Microsoft
- Duet (SIGCOMM 2014)
  - SLBs are bad for performance and cost!
  - Hybrid design using commodity switches and servers.
- Maglev (NSDI 2016)
  - Software based LB at Google
  - SLB performance is not too bad!

SilkRoad Overview

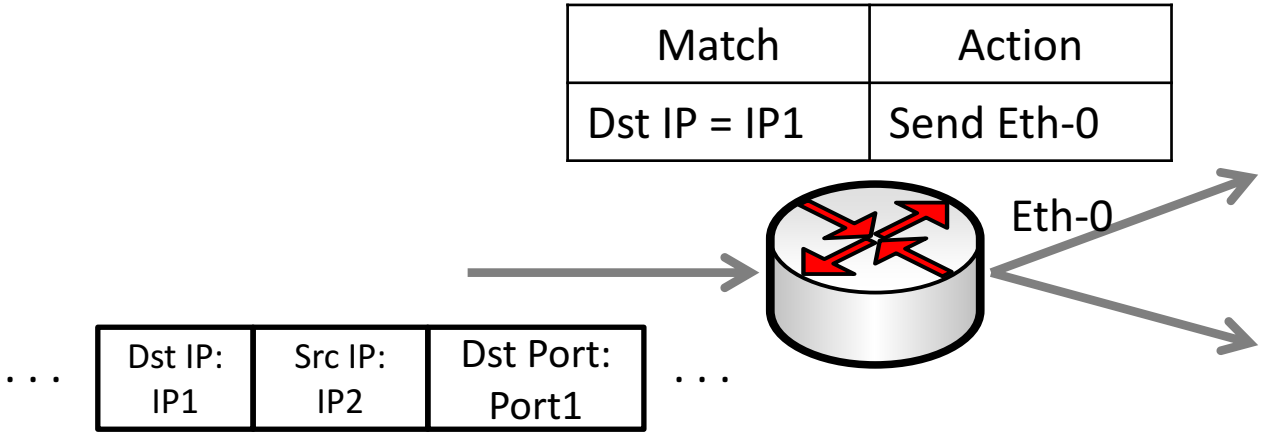
Load balancing entirely in hardware!

- Storing state:
  - Duet used small number of software instances
  - SilkRoad uses large memories on programmable switches

DRMT: Disaggregated Programmable Switching

Sharad Chole, Andy Fingerhut, and Sha Ma (Cisco Systems), Anirudh Sivaraman (Massachusetts Institute of Technology), Shay Vargaftik, Alon Berger, and Gal Mendelson (Technion), Mohammad Alizadeh (Massachusetts Institute of Technology), Shang-Tse Chuang (Cisco Systems), Isaac Keslassy (Technion, VMware), Ariel Orda (Technion), and Tom Edsall (Cisco Systems)

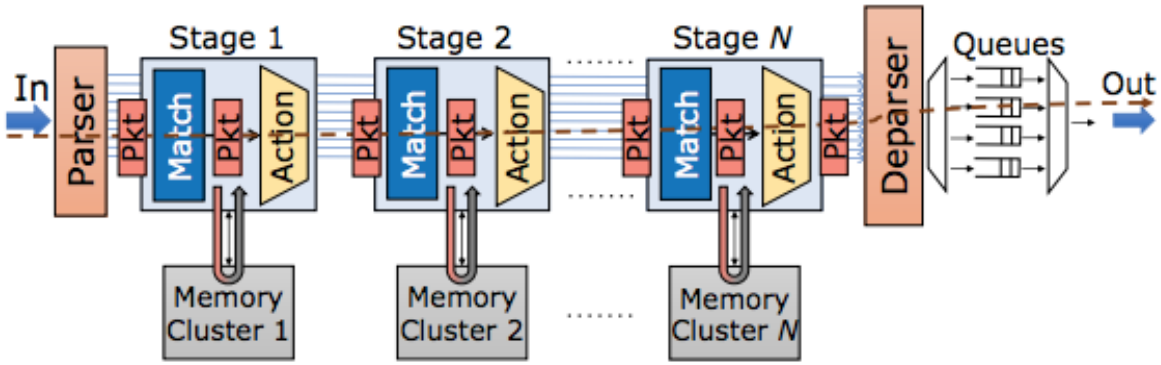
What Is Programmable Switching?



- Fixed-function switching:
  - Match/action on fixed packet header fields (e.g., dst IP, dst. MAC etc.)
- Programmable switching:
  - Match/action on arbitrary packet header bits
  - Programmed typically using P4.

How Is Programmable Switching Done Today?

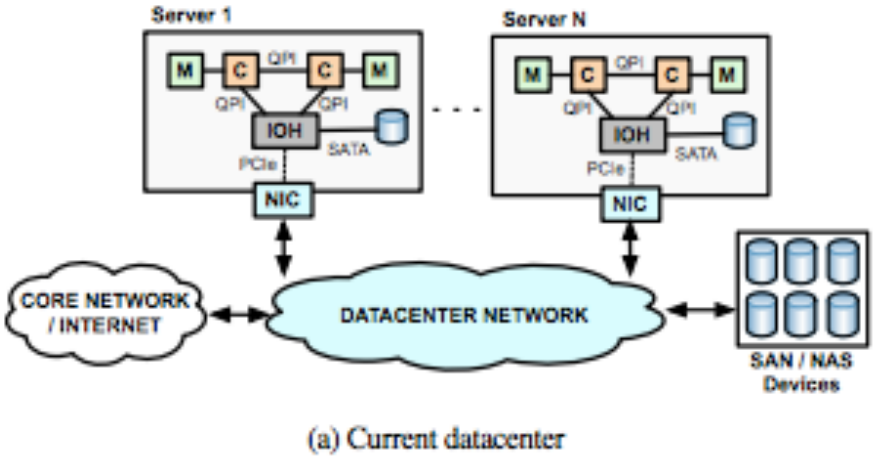
RMT[1] (Re-configurable Match Table)

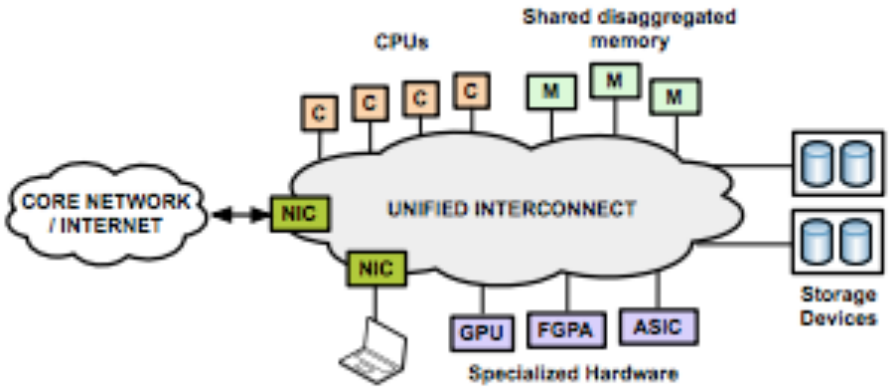


- Multiple logical stages:
  - Match/action on arbitrary bits
  - Variable depth
- Fixed number of physical stages:
  - Fixed processing and memory resources

[1]Forwarding Metamorphosis: Fast Programmable Match-Action Processing in Hardware for SDN, SIGCOMM 2013

Dis-aggregation in Datacenters

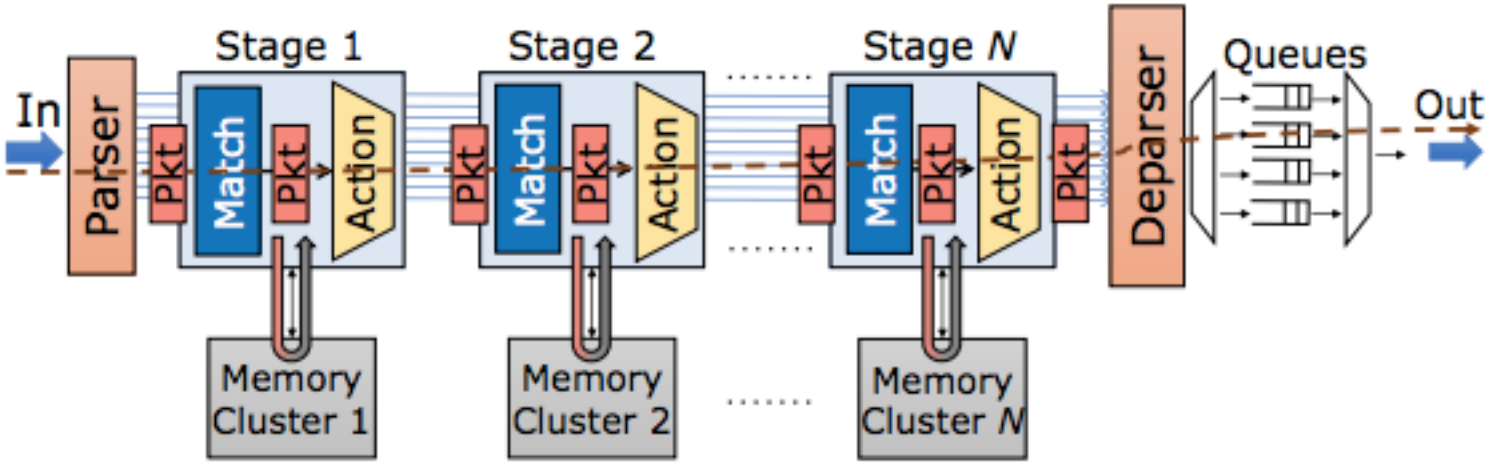




"Network requirements for resource disaggregation", OSDI 2016

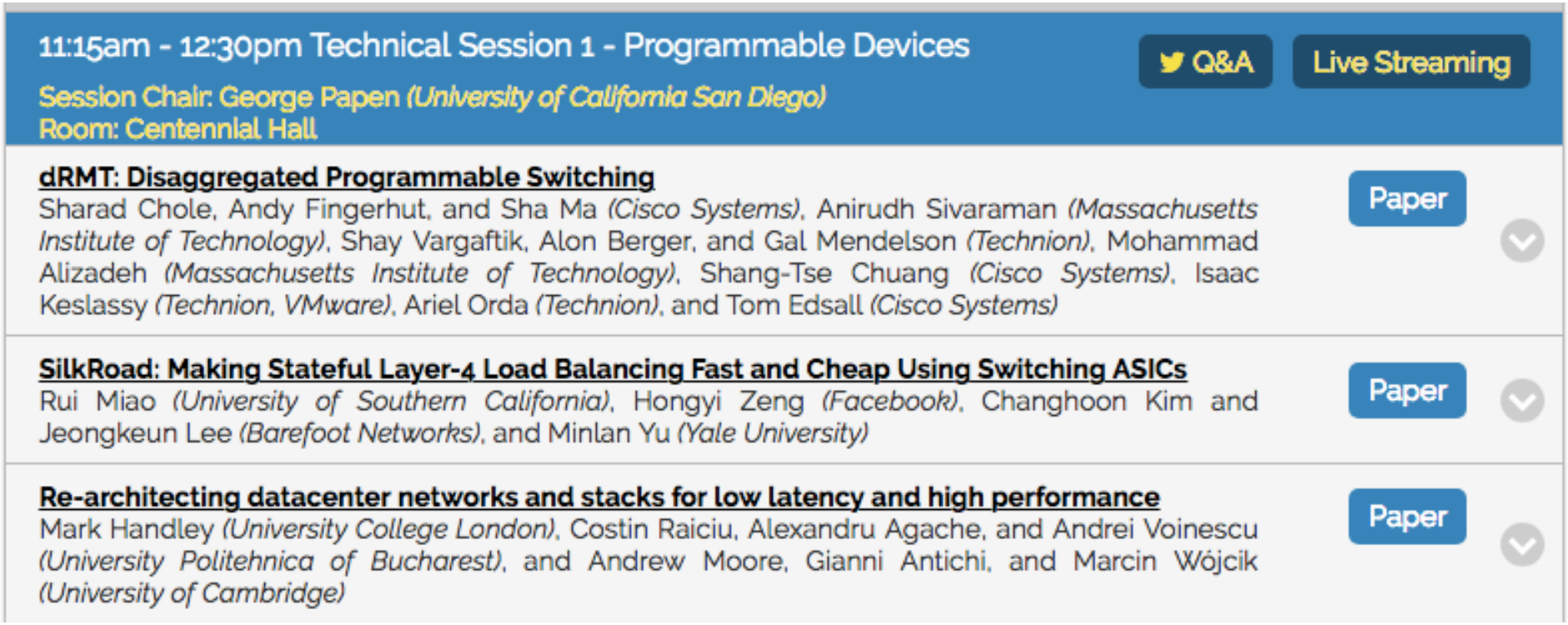
- Dis-aggregation helps improve: (1) resource utilization, (2) scalability, (3) time to adoption
- Existing solutions: Intel RSA, HP "The Machine", Facebook's Disaggregated Rack, Huawei's DC3.0, Berkeley FireBox, SeaMicro, soNUMA

dRMT (dis-aggregated RMT) Overview



- Problem: RMT architecture results in poor resource utilization
- Key Idea: Use resource dis-aggregation.
  - Memory: Decouple memory from individual stages
  - Compute: No packet movement across processors

Thank you!



11:15am - 12:30pm Technical Session 1 - Programmable Devices

Q&A Live Streaming

Session Chair: George Papen *(University of California San Diego)*
Room: Centennial Hall

dRMT: Disaggregated Programmable Switching
Sharad Chole, Andy Fingerhut, and Sha Ma *(Cisco Systems)*, Anirudh Sivaraman *(Massachusetts Institute of Technology)*, Shay Vargaftik, Alon Berger, and Gal Mendelson *(Technion)*, Mohammad Alizadeh *(Massachusetts Institute of Technology)*, Shang-Tse Chuang *(Cisco Systems)*, Isaac Keslassy *(Technion, VMware)*, Ariel Orda *(Technion)*, and Tom Edsall *(Cisco Systems)*

Paper

SilkRoad: Making Stateful Layer-4 Load Balancing Fast and Cheap Using Switching ASICs
Rui Miao *(University of Southern California)*, Hongyi Zeng *(Facebook)*, Changhoon Kim and Jeongkeun Lee *(Barefoot Networks)*, and Minlan Yu *(Yale University)*

Paper

Re-architecting datacenter networks and stacks for low latency and high performance
Mark Handley *(University College London)*, Costin Raiciu, Alexandru Agache, and Andrei Voinescu *(University Politehnica of Bucharest)*, and Andrew Moore, Gianni Antichi, and Marcin Wójcik *(University of Cambridge)*

Paper